

# Self-adaptive wavelet threshold denoising based on multi-resolution analysis

Xiaojun Yu\*

*School of Computer Engineering, Jiangsu University of Technology, Changzhou 213001, Jiangsu, China*

*Received 1 June 2014, www.cmnt.lv*

---

## Abstract

Various noises are usually mixed in the collection, processing or transmission of digital image, which reduces the image quality and which is bad for the subsequent image analysis; therefore, the image denoising processing is an essential link to conduct subsequent image analysis. With continuous development and improvements of wavelet theory, its excellent time-frequency characteristics have led to its extensive applications in image denoising. By analyzing the basic principle of wavelet threshold denoising, this paper has proposed a denoising algorithm of self-adaptive wavelet threshold. This algorithm decomposes and reconstructs the signal by using multi-resolution analysis; designs and constructs appropriate threshold function and realizes a new self-adaptive threshold denoising algorithm by optimizing the threshold with the threshold function. The experimental result demonstrates that compared with median filtering algorithm and mean filtering algorithm, the algorithm of this paper can improve the signal to noise ratio; maintain the detail information and texture features of the image over denoising and have better denoising effects.

*Keywords:* image denoising, wavelet analysis, multi-resolution

---

## 1 Introduction

The main media for people to transmit information are voice and image, however, image is usually interfered by various noises during its generation and transmission and the image quality will also be damaged, which is bad for the subsequent higher image processing [1]. Therefore, in the image pre-processing, it is very necessary to perform image denoising since it can improve the signal to noise ratio of the image and highlight the expected characteristics of the image. People have been pursuing a method which can not only effectively reduce noises, but also preserve the edge information of the image [2].

People have developed numerous denoising methods according to the image characteristics, rules and statistical characteristics of spectrum distribution of noise. Among these methods, the most intuitive one is to transform the noisy signal to the frequency domain with Fourier transform according to the feature that the noise energy usually focuses on the high frequency while the signal spectrum is distributed on finite interval and then perform filtering denoising by adopting low-pass filtering [3]. However, since the image details are also distributed in high-frequency regions, this method will also smooth the image edges in the image denoising and lose some detail information of the image. Worse still, the signal fantastic points with significant importance in signal detection may also be filtered [4]. Therefore, the denoising methods based on traditional Fourier transform have some contradictions in protecting the signal edge and suppressing noises and it is difficult to correctly identify and remove the noises in the signal. A dilemma in image

denoising is how to balance the denoising and the image details preservation [5].

Wavelet transform has good property of time-frequency localization, which provides an excellent tool to solve this problem. Multi-resolution analysis can be adopted to transform the noisy signal to the wavelet domain with wavelet transform, which can depict the non-stationary characteristics of the signal such as the edge, peak and breakpoint for ease of characteristic extraction. Besides, wavelet transform also has the characteristics of low entropy and de-correlation. Wavelet transform can de-correlate the signal and the noises have a whitening trend. Wavelet coefficients are sparse. Usually, the signal corresponds to few big wavelet coefficients while the noise to many small wavelet coefficients, which are good for signal denoising [6]. This paper firstly investigates the basic principle of wavelet transform threshold denoising. Then it analyzes the selection methods of wavelet coefficient threshold, the threshold estimation methods and the self-adaptive denoising algorithm. Finally, it realizes the self-adaptive filtering of wavelet transform based on the theoretical analysis and experiments.

## 2 Wavelet multi-resolution analysis and mallet algorithm

Image denoising is a classical problem in signal processing. The traditional denoising methods usually use mean or linear methods such as Wiener filtering, but the denoising effects are not very satisfactory. As wavelet theory is improved gradually, it has drawn an increasing number of eyes in image signal denoising with its excellent

---

\* *Corresponding author's* e-mail: 22288492@qq.com

time-frequency characteristics and it has also pioneered the denoising with non-linear methods.

The idea of wavelet transform is built upon the window function with automatically adjustable length. S.Mallat has raised the concept of multi-resolution analysis in the construction of orthogonal wavelet; given the mathematical explanation of orthogonal wavelet from the perspective of function analysis, illustrated the wavelet multi-resolution feature vividly in the concept of space. Proposed the common method for orthogonal wavelet construction; unified the common method with all the previous construction methods and proposed a fast algorithm of wavelet transform, namely Mallat algorithm by using the algorithm similar to fast Fourier transform in Fourier analysis [7].

Multi-resolution analysis is to construct a group of function space and the constitution of every group of space has a unified form, but the closure all space approximate  $L^2(R)$ . In every space, all the functions constitute the standard orthogonal basis of this space while the function of the closure of all function space form the standard orthogonal basis of  $L^2(R)$ . If decomposing the signal in such space, the mutually orthogonal time-frequency characteristics can be obtained. Additionally, because the number of the space is infinite and countable, it is convenient to analyze some characteristics of the signal we care about [8].

When using computer or other digital devices to realize wavelet transform,  $t$  should take discrete value and it is very necessary to investigate the wavelet transform with  $a, b$  and  $t$  discrete values and to develop a set of fast wavelet transform algorithm. The MRA theory founded by Mallat S has played a key role in the development of fast wavelet algorithm.

The basic idea of MRA theory is to observe the signal with different resolutions. When the scale changes, the signal can be observed roughly and elaborately.

$$\phi\left(\frac{t}{2^j}\right) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_0(k) \phi\left(\frac{t}{2^{j-1}} - k\right), \quad (1)$$

$$\psi\left(\frac{t}{2^j}\right) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_1(k) \phi\left(\frac{t}{2^{j-1}} - k\right), \quad (2)$$

where  $h_0(k)$  and  $h_1(k)$  are the weighting coefficients and Equations (1) and (2) are two-scale relation equations, which reveal the relationship between wavelet function and scale function in MRA.

Inspired by the pyramid encoding algorithm in the image processing by Burt and Adelson, Mallat realizes wavelet transform and inverse wavelet transform, namely the famous Mallat decomposition and reconstruction algorithms based on MRA theory[9].

Make  $h_0(k)$  and  $h_1(k)$  the two filters meeting Equations (1) and (2) and  $a_j(k)$  and  $d_j(k)$  are the discrete approximation coefficients in MRA. Then  $a_j(k)$  and  $d_j(k)$  meet the recurrence relation between Equations (3) and (4).

$$a_{j+1}(k) = \sum_{n=-\infty}^{\infty} a_j(n) h_0(n-2k) = a_j(k) * \bar{h}_0(2k), \quad (3)$$

$$d_{j+1}(k) = \sum_{n=-\infty}^{\infty} a_j(n) h_1(n-2k) = a_j(k) * \bar{h}_1(2k), \quad (4)$$

where  $\bar{h}(k) = h(-k)$  and get the stepwise decomposition of the coefficient from the recursion formula The recurrence relation between Equations (3) and (4) is the formula of Mallat decomposition algorithm and Equation (5) is the equation of Mallat reconstruction algorithm[10].

$$a_j(k) = \sum_{n=-\infty}^{\infty} a_{j+1}(n) h_0(k-2n) + \sum_{n=-\infty}^{\infty} d_{j+1}(n) h_1(k-2n). \quad (5)$$

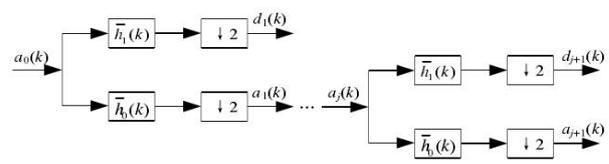


FIGURE 1 Wavelet decomposition

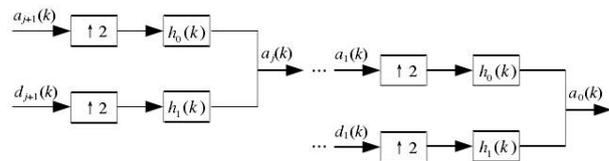


FIGURE 2 Wavelet reconstruction

Figures 1 and 2 are the sketches of wavelet decomposition and wavelet reconstruction respectively. When performing wavelet decomposition, overturn  $h_0(k)$  and  $h_1(k)$ , namely turn  $h_0(k)$  and  $h_1(k)$  into  $h_0(-k)$  and  $h_1(-k)$ , while in wavelet reconstruction, it doesn't need to overturn  $h_0(k)$  and  $h_1(k)$ . There are two decimations and two interpolations in the decomposition and reconstruction respectively and the filters used in decomposition and reconstruction are the same.

LL <sub>3</sub>	LH <sub>3</sub>	LH <sub>2</sub>	LH <sub>1</sub>
HL <sub>3</sub>	HH <sub>3</sub>	HL <sub>2</sub>	
HL <sub>1</sub>		HH <sub>2</sub>	HH <sub>1</sub>
HL <sub>1</sub>		HH <sub>1</sub>	

FIGURE 3 The sketch of image wavelet decomposition (three layers)

Figure 3 is the sketch of image wavelet three-layer decomposition. LL<sub>3</sub> is low-frequency component; LH<sub>1</sub>, LH<sub>2</sub> and LH<sub>3</sub> are the horizontal high-frequency components from the first layer to the third layer respectively; HL<sub>1</sub>, HL<sub>2</sub> and HL<sub>3</sub> are the vertical high-frequency components from the first layer to the third layer respectively and HH<sub>1</sub>, HH<sub>2</sub> and HH<sub>3</sub> are the diagonal high-frequency components from the first layer to the third layer. The low-frequency components preserve the general information of the original image while the high-frequency

components include the detail information of the image [11].

### 3 Self-adaptive wavelet threshold denoising algorithm

#### 3.1 THRESHOLD PROCESSING FUNCTION

Donoho has divided the threshold processing function into soft threshold function and hard threshold function. Many improved threshold functions have emerged with continuous improvements in threshold processing function and the improved soft-hard eclectic threshold processing function is as follows:

$$\hat{w}_{jk} = \begin{cases} \text{sign}(w_{j,k}) (|w_{j,k}| - a\lambda), & |w_{j,k}| \geq \lambda \\ 0, & |w_{j,k}| < \lambda \end{cases} \quad (0 \leq a \leq 1), \quad (6)$$

where sign (.) is sign function. In particular, when  $a$  takes 0 and 1, the above formula will become the hard threshold and soft threshold estimation method. As for  $0 < a < 1$ , the estimate data  $\hat{w}_{j,k}$  by this method is between the soft and hard threshold methods; therefore, it is called soft-hard eclectic threshold method [12].

The idea of this method is very simple and common, but it also has good denoising effects. It can be seen through careful analysis that because  $\hat{w}_{j,k}$  estimated by the pure soft threshold method has a absolute value  $\lambda$  smaller than  $w_{j,k}$  ( $w_{j,k} \geq \lambda$ ), it needs to reduce this difference with all efforts. The soft-hand eclectic threshold method meets this requirement. However, given practical consideration, the difference is not reduced to 0 because if reducing this difference to 0, this method will evolve as the hard threshold estimation method of wavelet coefficient and it can satisfy the requirements to improve the performance. In order to obtain better denoising effects, this paper has adopted a new threshold function based on the soft-hard eclectic threshold method and its formula is as follows:

$$\hat{w}_{j,k} = \begin{cases} \text{sign}(w_{j,k}) \mu (|w_{j,k}| - a\lambda) & (0 \leq a \leq 1), \\ 0 \end{cases} \quad (7)$$

and:

$$\mu = \begin{cases} 1, & |w_{j,k}| \geq \lambda \\ \frac{\lambda}{\lambda - \lambda_0}, & \lambda_0 \leq |w_{j,k}| < \lambda \\ 0, & |w_{j,k}| < \lambda_0 \end{cases}, \quad (8)$$

where, sign (.) is sign function and it can adjust the soft and hard threshold functions by introducing weight function  $\mu$ , namely it can prevent the useful information from being filtered as noises while preserve the edge information; thus, it can further improve the quality of the reconstructed image.  $\lambda_0$  is the undetermined parameter and it is usually  $\delta$  times of  $\lambda$ , namely  $\lambda_0 = \delta\lambda$ . After repeated

experiments and comparison,  $\lambda_0$  is 0.6 times of  $\lambda$  in order to improve the quality of the reconstructed image.

It can be seen from the linear nature of wavelet transform,  $w_{j,k}$  is made up of  $u_{j,k}$  and  $v_{j,k}$  probably because of the influence of  $v_{j,k}$  (for most  $w_{j,k}$ ) and its purpose is to minimize  $\|\hat{w}_{j,k} - u_{j,k}\|$ . Therefore, to make the value of  $|\hat{w}_{j,k}|$  between  $|w_{j,k}| - \lambda$  and  $|w_{j,k}|$  may make the estimated wavelet coefficient  $\hat{w}_{j,k}$  more approximate to  $u_{j,k}$ . Based on this idea, add factor  $a$  in the threshold estimator and properly adjust  $a$  between 0 and 1 to get better denoising effects. When  $a$  takes 0 and 1, the above formula will become the hard and soft threshold estimation method. Properly adjust  $a$  between 0 and 1 and get better denoising effects [13,14].

#### 3.2 THE THRESHOLD OF THIS PAPER

The selection of threshold directly affects the denoising result. If the threshold is too big, many wavelet coefficients will be set as 0 and it will damage many signal details; if the threshold is too small, it fails to reach the expected effects. There are many threshold estimations algorithms and the commonly-used one is the fixed threshold with a threshold of  $\lambda = \sigma\sqrt{2\ln N}$ , where  $N$  is the number of total wavelet decomposition coefficients in all scales of the noisy signal and  $\sigma$  is the noise variance. In practical application, the variance of noise signal  $\sigma$  is usually unknown and it can be estimated as follows by the first-layer wavelet decomposition coefficient:

$$\sigma = MAD/0.6745, \quad (9)$$

where  $MAD$  is the median value of the absolute value of the first-layer wavelet decomposition coefficient and 0.6745 is the adjustment coefficient of the standard deviation of Gaussian noise. The noise signal variance estimated by the above formula meets the condition of  $\sigma < \hat{\sigma} \leq 1.01\sigma$  and the noise variance  $\sigma$  can be approximately estimated from  $\sigma$ .

In the threshold filtering, if the threshold  $\lambda$  is too small, there are too many signal losses; if  $\lambda$  is too big, it can preserve many details and it has unclear denoising and affects the filtering effects. It can be known from the noise wavelet transform characteristics that the wavelet coefficient after noise wavelet transform has the mean of 0 and the variance of white noise. It increases with the increase of the scale  $j$  and the wavelet coefficient of white noise will reduce. The zero-mean Gaussian white noise is the Gaussian noise with the orthogonal wavelet transform as zero mean and the variance of the noise coefficient in different scales decreases with the increase of the scale, namely that Gaussian white noise has negative singularity. Gaussian white noise is the distribution of unified

Lipschitz Exponent  $\alpha < 0 (\alpha = -\frac{1}{2} - \varepsilon, \varepsilon > 0)$ . The discrete white noise is almost singular everywhere and in

the multi-scale wavelet analysis, the wavelet spectrum of white noise gradually fades away with the increase of scale. However, the wavelet transform of effective signal still has clear manifestation in large scale. It can be seen from the characteristics of the above noise wavelet transform that the wavelet coefficient of white noise decreases with the increase of the scale; therefore, when using threshold method, it is unreasonable to use of fixed value for different scales and the threshold reduces correspondingly with the increase of the scale. The following improved formula will be adopted in this paper:

$$\lambda_j = \sigma \sqrt{2 \log N} / \ln(e + j - 1), \quad (10)$$

obviously, when  $j=1$ ,

$$\lambda_j = \sigma \sqrt{2 \log N} = \lambda. \quad (11)$$

Namely, the result is consistent with the result using the original threshold calculation formula when the scale  $j=1$ ; however, when the scale  $j>1$ , the threshold  $\lambda$  reduces with the increase of the scale, which is exactly consistent with the fact that the noise reduces gradually with the increase of the scale. Therefore, the improved formula is more reasonable.

In order to guarantee the smoothness of the new threshold function and the continuity of multi-order derivative and get close to the hard and soft threshold functions, the threshold function can be constructed as

$$\eta(x, \lambda) = \begin{cases} x + \lambda - \frac{\lambda}{2\beta + 1} & x \leq -\lambda \\ \frac{x^{2\beta + 1}}{(2\beta + 1)\lambda^{2\beta}} & |x| < \lambda \\ x - \lambda + \frac{\lambda}{2\beta + 1} & x > \lambda \end{cases}, \quad (12)$$

where,  $\beta$  is the selection parameter of the threshold function. It can be seen that when  $\beta = 0$ , this function approximates to the hard threshold function and when  $\beta$  goes to infinity, this function approximates to the soft threshold function.

### 3.3 SELF-ADAPTIVE WAVELET DENOISING ALGORITHM

Because the noises are mainly distributed in the high-frequency area, it needs only to process the high-frequency wavelet coefficients and the threshold self-adaptive wavelet denoising algorithm proposed by this paper based on multi-resolution analysis is as follows:

Perform  $J$  layer discrete wavelet transform (DWT) to the noisy image  $X$  and record the wavelet coefficient vector as vector  $Y$ :

$$Y = [A_J, H_J, V_J, D_J, H_{J-1}, V_{J-1}, D_{J-1}, \dots, H_1, V_1, D_1]$$

where, vector  $A$  is low-frequency coefficient; vector  $H$  is horizontal high-frequency coefficient; vector  $V$  is vertical high-frequency coefficient and vector  $D$  is diagonal high-frequency coefficient.

Get the lengths of every wavelet coefficient and store them in vector  $L$ . And  $L(1)$  is the length of vector  $A$ ;  $L(i)$  is the sum of the lengths of high-frequency coefficients ( $H, V, D$ ) ( $i=2,3,\dots,J+1$ ) and  $L(J+2, :)$  is the length of the entire  $X$ .

Perform threshold optimization and processing on the high-frequency components of the wavelet coefficient in different scales.

- 1) Initialize the threshold in different scales  $\lambda_j = \sqrt{2 \ln(L(j)) / L(j)}$  and the step length is  $u=0.001$ .
- 2) Take the high-frequency coefficients in different scales and store them in vector  $C_j$ ,  $C_j=[H_j, V_j, D_j]$
- 3) While  $(abs(\nabla \lambda_j(k)) / \lambda_j(k)) > 1e-6$  Calculate  $\Delta \lambda_j(k)$  and  $\lambda_j(k)$  respectively according to Formula (10);
- 4) Process the threshold  $\lambda_j(k)$  determined by Formula (10) in  $C_j$  according to Formula (12) and put the processed data into  $C_j$ .

Reconstruct the image through inverse wavelet transform by using the low-frequency components  $A_J$ , and the high-frequency wavelet coefficients

## 4 Experimental results and analysis

In order to test the performance of the algorithm of this paper, add Gaussian noise with  $\mu=0$  and a noise variance  $\sigma^2$  between 0.01~0.04 the 256x256 original grayscale image. Perform denoising on the image with the traditional 3x3 mean filtering, 3x3 median filtering and the algorithm of this paper respectively and compare the filtering effects by comparing their signal to noise ratio improvement factor  $R$  and peak signal to noise ratio  $PSNR$ .

It can be seen from Figure 4 that after adding high-intensity Gaussian noise to the original image, the image filtered by the algorithm of this paper has improved visual effects, clear edge and image details and improved image quality and it also has outstanding advantages compared with the traditional Gaussian noise filtering methods.

The experimental results are indicated in Tables 1 and 2. It can be seen from these two tables that the algorithm of this paper has bigger advantages compared with the traditional algorithms, as evidenced by the fact that the image signal to noise ratio has been enhanced greatly and the image quality has also been improve.

TABLE 1 Comparison of  $R$  of different filters on original image

Noise variance	Algorithm of this Paper	Mean denoising	Median denoising
0.010	-5.6371	-5.3627	-5.9877
0.015	-6.8978	-6.5273	-6.6462
0.020	-7.5723	-6.9362	-6.6372
0.025	-8.3536	-7.5721	-6.6852
0.030	-8.7387	-7.6273	-6.8126
0.035	-9.0152	-7.8516	-6.9273
0.040	-9.5173	-8.1553	-7.2763

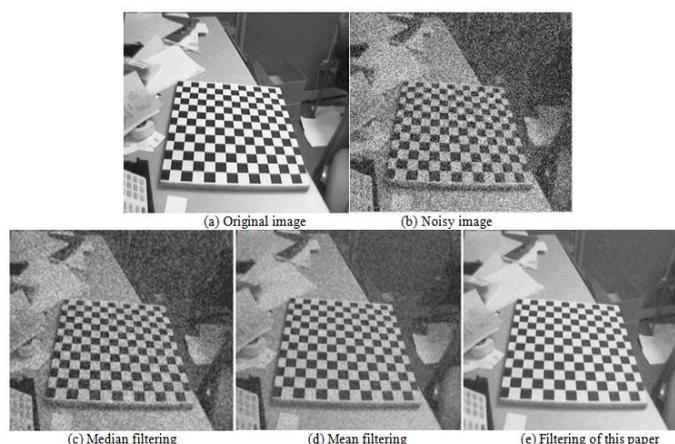
FIGURE 4 Comparison of denoising effect on original image with noise of  $\sigma^2=0.01$ 

TABLE 2 Comparison of PSNR of different filters on original image

Noise variance	Algorithm of this paper	Mean denoising	Median denoising	Noised image
0.010	27.3682	27.4163	27.3484	18.9335
0.015	26.8738	26.6472	26.4728	18.0163
0.020	25.4737	24.7621	24.3725	17.6372
0.025	24.8362	23.1036	23.0126	16.8738
0.030	23.7653	22.8377	21.9881	15.9356
0.035	23.1263	22.0265	21.1723	15.1362
0.040	22.8987	21.5811	20.5221	14.3726

## 5 Conclusions

By investigating the basic principle of wavelet transform threshold denoising, including the threshold selection method and the wavelet coefficient threshold estimation methods, this paper has proposed a self-adaptive denoising algorithm based on wavelet transform, realizes the self-adaptive filtering of wavelet transform based on the experiments and theoretical analysis and reduces the influence the subjective factors have on denoising effects. The simulation experiment shows that this algorithm protects the image details and has better denoising effects while it has effective filtering.

## References

- [1] Jaiswal A, Upadhyay J, Somkuwar A 2014 Image Denoising and Quality Measurements by Using Filtering and Wavelet Based Techniques *AEU – International Journal of Electronics and Communications* **68**(8) 699-705
- [2] Mukhopadhyay S, Mandal J K 2014 Wavelet based Denoising of Medical Images Using Sub-band Adaptive Thresholding through Genetic Algorithm *Procedia Technology* **10** 680-9
- [3] Gradolewski D, Redlarski G 2014 Wavelet-based Denoising Method for Real Phonocardiography Signal Recorded by Mobile Devices in Noisy Environment *Computers in Biology and Medicine* **52**(1) 119-29
- [4] Zhao Y 2013 Multi-level Denoising and Enhancement Method Based on Wavelet Transform for Mine Monitoring. *International Journal of Mining Science and Technology* **23**(1) 163-6
- [5] Hill P R, Achim A M, Bull D R, Al-Mualla M E 2014 Dual-tree Complex Wavelet Coefficient Magnitude Modelling Using The Bivariate Cauchy-Rayleigh Distribution for Image Denoising *Signal Processing* **105** 464-72
- [6] Sun H, Zi Y, He Z 2014 Wind Turbine Fault Detection Using Multiwavelet Denoising with The Data-Driven Block Threshold *Applied Acoustics* **77** 122-9
- [7] Haltmeier M, Munk A 2014 Extreme Value Analysis of Empirical Frame Coefficients and Implications for Denoising by Soft-Thresholding *Applied and Computational Harmonic Analysis* **36**(3) 434-60
- [8] Le Pogam A, Hanzouli H, Hatt, Le Rest C C, Visvikis D 2013 Denoising of PET Images by Combining Wavelets and Curvelets for Improved Preservation of Resolution and Quantitation *Medical Image Analysis* **17**(8) 877-91
- [9] Mert A, Akan A 2014 Detrended Fluctuation Thresholding for Empirical Mode Decomposition Based Denoising *Digital Signal Processing* **32** 48-56
- [10] Paul R, Sengupta A, Pathak R R 2013 Wavelet Based Denoising Technique for Liquid Level System *Measurement* **46**(6) 1979-94
- [11] Om H, Biswas M 2014 MMSE Based Map Estimation for Image Denoising *Optics & Laser Technology* **57** 252-64
- [12] Hoseini M R, Zuo M J, Wang X 2014 Denoising Ultrasonic Pulse-Echo Signal Using Two-Dimensional Analytic Wavelet Thresholding *Measurement* **45**(3) 255-67
- [13] Zhou Y, Zhao H, Shang L, Liu T 2014 Immune K-SVD Algorithm for Dictionary Learning in Speech Denoising *Neurocomputing* **137**(5) 223-33
- [14] Kabir A, Shahnaz C 2012 Denoising of ECG Signals Based on Noise Reduction Algorithms in EMD and Wavelet Domains *Biomedical Signal Processing and Control* **7**(5) 481-9

## Author



Xiaojun Yu, January 1979, China.

**Current position, grades:** researcher Jiangsu University of Technology, China.

**University studies:** bachelor degree at Liaoning Technical University. Master degree at Liaoning Technical University in 2005.

**Scientific interests:** vector machine and data mining.